**Assignment: Logistic Regression**

**521H0489 – Hồ Hữu An**

# **Loss Function using Cross Entropy:**

Loss=−(y ⋅ log(y^) + (1-y).log(1-y^)

y is the actual label (0 or 1).

y^ is the predicted probability that y=1 given the input features.

# **Calculating Derivatives:**

To update the parameters of the model, you'll need to compute the gradients of the loss function with respect to each parameter (weights and bias).

Assuming a simple logistic regression with one feature and a bias term, the gradients can be calculated as follows:

For weights:

For bias:

x is the input feature.

* **Chain Rule**

If z = f(y) and y = g(x) or z = f(g(x))

Then

Application of calculating the derivative of the sigmoid:

# **Updating Parameters using Gradient Descent:**

For each parameter p (weight or bias), the update rule for Gradient Descent can be defined as:

is the learning rate, a hyperparameter that determines the step size in the parameter space.

# **Logistic Regression for Binary Classification:**

In Logistic Regression, the model predicts the probability that an input belongs to a certain class (e.g., Class 0 or Class 1). The logistic function (also known as the sigmoid function) is used to map predictions between 0 and 1.

Linear Model: z = b + w1 x1 + w2 x2 + . . . + wn xn

Logistic (Sigmoid) Function:

Learning Process:

1. Initialize Parameters: Start by initializing the weights (w1, w2, w3, …, wn ) and the bias term (b) with small random values.
2. Forward Propagation:
   1. Compute the weighted sum of inputs and the bias: ​ z = b + w1 x1 + w2 x2 + . . . + wn xn
   2. Pass the result through the logistic function to get the predicted probability:
3. Calculate Loss (Cross Entropy Loss):
   1. Compute the loss between the predicted probability (y^) and the true label (y).
   2. Cross Entropy Loss:

Loss=−(y ⋅ log(y^) + (1-y).log(1-y^)

1. Compute Gradients:

Calculate the gradients of the loss function with respect to each parameter:

For weights:

For bias:

1. Update Parameters using Gradient Descent:

Update the weights and bias by subtracting the gradient scaled by the learning rate (

1. Repeat: Repeat steps 2 to 5 for multiple iterations (epochs) until the model converges or a stopping criterion is met.
2. Prediction:

After training, use the learned parameters to make predictions on new data by passing it through the logistic function.